

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-067696

(43)Date of publication of application : 11.03.1994

(51)Int.Cl.

G10L 9/08
G10L 9/18

(21)Application number : 04-244038

(71)Applicant : SONY CORP

(22)Date of filing : 21.08.1992

(72)Inventor : MATSUMOTO ATSUSHI
KATAYANAGI KEIICHI
NISHIGUCHI MASAYUKI

(54) SPEECH ENCODING METHOD

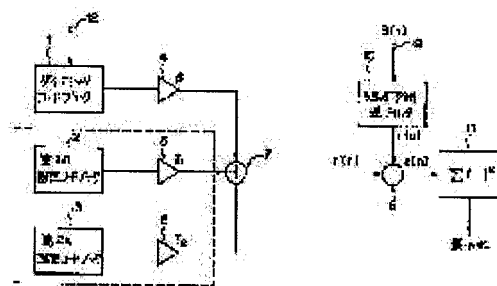
(57)Abstract:

PURPOSE: To reduce the arithmetic operation quantity by detecting the analytic output of a past speech signal which has the shortest distance to the analytic output of a current input speech signal and encoding the current input speech on the basis of this detection output.

CONSTITUTION: A speech signal $S(n)$ which is sampled, for example, at a 8kHz sampling frequency f_s and converted into a digital signal is inputted to an input terminal 9. A short-period predictive inverse filter 10 filters the input signal reversely to predict a sound

generated at the inner part of the throat, and its residual output $r(n)$ is supplied to a subtracter 8; and the output of the subtracter 8 is supplied as an output difference $e(n)$ to an energy calculation part ($\Sigma()$ 2) 11. Then this

energy calculation part 11 calculates the energy of the output difference $e(n)$ between the residual output $r'(n)$ and the residual output $r(n)$ from a short-period predictive inverse filter 10 and then selects, for example, the dynamic code vector of a dynamic code vector 1 from a terminal 12 so as to minimize the energy of the output difference $e(n)$.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)特許出願公開番号

特開平6-67696

(43)公開日 平成6年(1994)3月11日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/08	G	8946-5H		
	J	8946-5H		
9/18	E	8946-5H		

審査請求 未請求 請求項の数3(全 10 頁)

(21)出願番号 特願平4-244038

(22)出願日 平成4年(1992)8月21日

(71)出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72)発明者 松本 淳

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72)発明者 片柳 恵一

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72)発明者 西口 正之

東京都品川区北品川6丁目7番35号 ソニー株式会社内

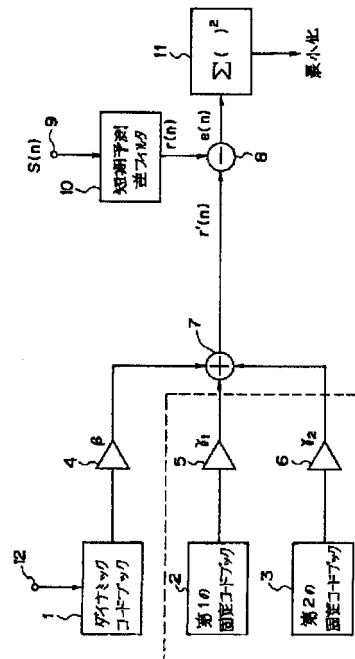
(74)代理人 弁理士 小池 晃 (外2名)

(54)【発明の名称】 音声符号化方法

(57)【要約】

【構成】 短期予測逆フィルタ10は入力端子9から供給された音声信号 $S(n)$ を残差信号 $r(n)$ とする。減算器8はダイナミックコードブック1等からのダイナミックコードベクトル等により合成された残差信号 $r'(n)$ と残差信号 $r(n)$ との差を $e(n)$ として出力する。エネルギー計算部11は、出力差 $e(n)$ のエネルギーを計算する。そして、このエネルギーを最小とするようなダイナミックコードベクトルのインデックスを端子12からサーチし、取り出すことによって音声符号化が行われる。

【効果】 演算量を大幅に減らせる。



1

【特許請求の範囲】

【請求項1】 過去の音声信号との相関を利用して現在の入力音声信号を分析し符号化する音声符号化方法において、

上記現在の入力音声信号を分析する工程と、

この入力音声信号の分析出力との距離が最短となる上記過去の音声信号の分析出力を検出する検出工程と、

上記検出工程の検出出力を基に現在の入力音声信号を符号化する符号化工程とを有することを特徴とする音声符号化方法。

【請求項2】 過去の音声信号の分析出力が複数のコードベクトルとして記憶されたコードブックをサーチして現在の入力音声信号との相関を利用して符号化を行う符号化方法であって、

現在の入力音声信号を分析する工程と、

この入力音声信号の分析出力との距離が最短となる上記コードブック内のコードベクトルを直接サーチする工程と、

上記サーチされて得られたコードベクトルのインデックスを用いて現在の入力音声信号を符号化する工程とを有することを特徴とする音声符号化方法。

【請求項3】 過去の音声信号の分析出力が複数のコードベクトルとして記憶されたコードブックをサーチして現在の入力音声信号との相関を利用して符号化を行う符号化方法であって、

現在の入力音声信号を分析する工程と、

この入力音声信号の分析出力との距離が最短となる上記コードブック内のコードベクトルを直接サーチする第1のサーチ工程と、

上記第1のサーチ工程で得られたコードベクトルを含め、該コードベクトルの近傍のコードベクトルの中から入力音声信号との相関が最適となるようなコードベクトルを間接サーチする第2のサーチ工程と、

上記第2のサーチ工程で得られたコードベクトルのインデックスを用いて現在の入力音声信号を符号化する工程とを有することを特徴とする音声符号化方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、低ビットレートの音声符号化方法に関し、過去の音声信号との相関を利用して現在の入力音声信号を分析し符号化する音声符号化方法に関する。

【0002】

【従来の技術】近年、低ビットレート、すなわち、4.8～9.6 kbpsでの音声符号化方法には、ベクトル和励起ニア予測（VSELP: Vector Sum Excited Linear Prediction）等のコード励起ニア予測（CELP: Code Excited Linear Prediction）が提唱されている。

【0003】このVSELPについての技術内容は、モトローラ・インコーポレーテッドによる特表平2-50

2

2135号公報の「改良されたベクトル励起源を有するデジタル音声コーダ」及び「VECTOR SUM EXCITED LINEAR PREDICTION(VSELP): SPEECH CODING AT 8 KBPS: Ira A. Gerson and Jasiuk: Paper presented at the Int. Conf. on Acoustics, Speech and Signal Processing - April 1990」に記載されている。

【0004】このVSELPを用いた音声符号化方法は、アナリシスバイシンセシス（Analysis by synthesis）によるコードブックサーチにより、低ビットレートによる高品質音声伝送を実現している。また、VSELPを用いた音声符号化方法を適用した音声符号化装置（音声コーダ）においては、入力音声信号の特性を形成するピッチ及びノイズ予測器が導入された時間変動線形フィルタをコードブックに記憶されたコードベクトルを選択することで励起させて音声信号を符号化している。

【0005】具体的には、音声の各フレームに対して、音声コーダはそれぞれのコードベクトルをフィルタに導入して音声信号を形成し、もとの入力音声信号と比較してその差分出力を検出する。この差分出力は人間の聴覚に基づく重み付フィルタに通され重み付けされる。ここで、現在のフレームに対して最小のエネルギーで重み付けされた差分出力を発生させるコードベクトルを選択することが望まれる。

【0006】図5は上述したようなVSELPを用いた音声符号化方法を適用した音声コーダの概略構成を示す機能ブロック図である。この音声コーダは、過去の音声信号の分析出力が複数のコードベクトルとして記憶されたダイナミックコードブック51と、雑音成分に関したコードベクトルを記憶する第1の固定コードブック52及び第2の固定コードブック53の合わせて3つのコードブックを有する。

【0007】入力端子60には例えばサンプリング周波数 $f_s = 8 \text{ kHz}$ でサンプリングされ、図示しないバンドパスフィルタ（BPF）によって帯域が選択され、A/D変換器によってデジタル信号に変換された音声信号 $S(n)$ が入力される。この音声信号 $S(n)$ は減算器59に供給される。一方、この減算器59には上記3つのコードブックから選択された各コードベクトルで合成された音声 $S'(n)$ も供給される。この減算器59の出力は出力差 $e(n)$ としてエネルギー計算部61（ $\Sigma(\quad)^2$ ）に供給される。このエネルギー計算部61は、上記出力差 $e(n)$ のエネルギーを計算する。そして、その出力差 $e(n)$ のエネルギーを最小とするように、例えば端子50からダイナミックコードブック51のダイナミックコードベクトルが選択される。同様に上記第1の固定コードブック52及び第2の固定コードブック53についても出力差 $e(n)$ のエネルギーが最小となるような第1の固定コードベクトル及び第2の固定コードベクトルが選択される。

【0008】つまり、上記3つのコードブックから各コ

10

20

30

40

50

ードベクトルをサーチするには、上記ダイナミックコードブック51に記憶されているダイナミックコードベクトル、上記第1の固定コードブック52に記憶されている第1の固定コードベクトル及び上記第2の固定コードブック53に記憶されている第2の固定コードベクトルが合成されることにより形成された音声 $S'(n)$ と入力音声信号 $S(n)$ との出力差 $e(n)$ のエネルギーが最小となることが条件となる。

【0009】まず、上記ダイナミックコードブック51に記憶されているダイナミックコードベクトルの選択について以下に述べる。

【0010】上記ダイナミックコードブック51は、例えばサンプリング周波数を8kHzとし、音声の1フレームが40サンプルから構成されているとしたとき、例えば128個(128通り)のダイナミックコードベクトルを有する。この場合、上記合成フィルタ58は、128個のコードベクトルに演算を施す。そして、この合成フィルタ58から出力された音声 $S'(n)$ と入力音声 $S(n)$ との出力差 $e(n)$ を減算器59が算出し、エネルギー計算部61に供給する。このエネルギー計算部61は、上記出力差 $e(n)$ のエネルギーの値を算出する。そして、その出力差 $e(n)$ のエネルギーを最小とするように、端子50から最適なダイナミックコードベクトルを限定する最適インデックス j_{opt} をサーチする。

【0011】ここで、上記合成フィルタ58は、一般的に10次のIIRフィルタにより構成されているので1サンプルのデータに対して20回の積和演算が行われる。そのため、20回の積和演算を例えば40サンプルデータ分行い、さらにそれをコードベクトル128個分繰り返すことになる。

【0012】ここで、上記ダイナミックコードブック51のコードベクトルは、乗算器54で係数 β が乗算されたあと加算器57に供給される。また、上記第1の固定コードブック52のコードベクトルは、乗算器55で係数 γ_1 が乗算されたあと上記加算器57に供給され、また、第2の固定コードブック53のコードベクトルは、乗算器56で係数 γ_2 が乗算されたあと上記加算器57に供給される。上記加算器57は、上記それぞれの乗算器54、55及び56からの乗算結果を加算し、その加算出力を上記合成フィルタ58に供給している。

【0013】図5において、破線で囲んだ第1の固定コードブック52及び第2の固定コードブック53の各コードベクトルについても、上述した演算が上記合成フィルタ58で施される。なお、この場合、それぞれのコードベクトルの数は64通りであり、繰り返される演算の回数は64回となる。

【0014】

【発明が解決しようとする課題】ところで、上述のVSELPを用いた音声符号化方法を適用した音声コーデ

は、上述したようにコードブックのそれぞれのコードベクトルについてフィルタリングを全て行うと演算量が非常に多くなる。すなわち、ダイナミックコードブックをサーチする範囲の全ベクトルに対して、フィルタリングを行ったものについて原音声との距離を計算する必要があるその演算量は解析的な方法に比較してかなり多くなっている。VSELPによる音声符号化方法全体の処理時間の中でもこの部分の占める割合は他の部分に比較してかなり大きい。

【0015】本発明は、上記実情に鑑みてなされたものであり、演算量を低減できる音声符号化方法の提供を目的とする。

【0016】

【課題を解決するための手段】本発明に係る音声符号化方法は、過去の音声信号との相関を利用して現在の入力音声信号を分析し符号化する音声符号化方法において、上記現在の入力音声信号を分析する工程と、この入力音声信号の分析出力との距離が最短となる上記過去の音声信号の分析出力を検出する検出工程と、上記検出工程の検出出力を基に現在の入力音声信号を符号化する符号化工程とを有することを特徴として上記課題を解決する。

【0017】また、他の発明に係る音声符号化方法は、過去の音声信号の分析出力が複数のコードベクトルとして記憶されたコードブックをサーチして現在の入力音声信号との相関を利用して符号化を行う符号化方法であって、現在の入力音声信号を分析する工程と、この入力音声信号の分析出力との距離が最短となる上記コードブック内のコードベクトルを直接サーチする工程と、上記サーチされて得られたコードベクトルのインデックスを用いて現在の入力音声信号を符号化する工程とを有することを特徴として上記課題を解決する。

【0018】さらに、他の発明に係る音声符号化方法は、過去の音声信号の分析出力が複数のコードベクトルとして記憶されたコードブックをサーチして現在の入力音声信号との相関を利用して符号化を行う符号化方法であって、現在の入力音声信号を分析する工程と、この入力音声信号の分析出力との距離が最短となる上記コードブック内のコードベクトルを直接サーチする第1のサーチ工程と、上記第1のサーチ工程で得られたコードベクトルを含め、該コードベクトルの近傍のコードベクトルの中から入力音声信号との相関が最適となるようなコードベクトルを間接サーチする第2のサーチ工程と、上記第2のサーチ工程で得られたコードベクトルのインデックスを用いて現在の入力音声信号を符号化する工程とを有することを特徴として上記課題を解決する。

【0019】

【作用】現在の入力音声信号の分析出力との距離が最短となる過去の音声信号の分析出力を検出し、この検出出力を基に現在の入力音声信号を符号化するので、演算量を低減できる。

【0020】

【実施例】以下、本発明に係る音声符号化方法の実施例について、図面を参照しながら説明する。図1は本発明の音声符号化方法が適用された第1の実施例となる符号化装置（音声コーダ）の概略構成を示すブロック図である。

【0021】図1において、この音声コーダは、過去の音声信号の分析出力が複数のコードベクトルとして記憶されたダイナミックコードブック1と、雑音成分に関したコードベクトルを記憶する第1の固定コードブック2及び第2の固定コードブック3の合わせて3つのコードブックを有する。上記ダイナミックコードブック1は、時間と共に変化していくコードブックであり、一定時間過去の分析データに基づいて複数のコードベクトルが蓄えられている。また、第1の固定コードブック2及び第2の固定コードブック3には、雑音成分が固定された状態で蓄えられている。

【0022】入力端子9には例えばサンプリング周波数 $f_s = 8\text{ kHz}$ でサンプリングされ、図示しないA/D変換器によってデジタル信号に変換された音声信号 $S(n)$ が入力される。この音声信号 $S(n)$ は上述した従来例の合成フィルタ58の逆特性を有する短期予測逆フィルタ10に供給される。

【0023】この短期予測逆フィルタ10は、入力音声に逆フィルタをかけ、いわゆる喉の奥の部分の音を予測している。この短期予測逆フィルタ10からの出力を残差出力 $r(n)$ とする。この残差出力 $r(n)$ は、減算器8に供給される。

【0024】ここで、上記ダイナミックコードブック1のコードベクトルは、乗算器4で係数 β が乗算されたあと加算器7に供給される。また、上記第1の固定コードブック2のコードベクトルは、乗算器5で係数 γ_1 が乗算されたあと上記加算器7に供給される。また、第2の固定コードブック3のコードベクトルは、乗算器6で係数 γ_2 が乗算されたあと上記加算器7に供給される。そして、上記加算器7からは、残差出力 $r'(n)$ が上記減算器8に供給される。

【0025】すなわち、この減算器8には上記3つのコードブックの各コードベクトルに応じた残差出力 $r'(n)$ も供給される。この減算器8の出力は、出力差 $e(n)$ としてエネルギー計算部(Σ()²)11に供給される。

【0026】このエネルギー計算部11は、上記残差出力 $r'(n)$ と上記短期予測逆フィルタ10からの残差出力 $r(n)$ との出力差 $e(n)$ のエネルギーを計算する。そして、その出力差 $e(n)$ のエネルギーを最小とするように、この第1の実施例は、例えば端子12からダイナミックコードブック1のダイナミックコードベクトルを選択する。

【0027】つまり、上記ダイナミックコードブック1

からダイナミックコードベクトルをサーチするには、上記ダイナミックコードブック1に記憶されているダイナミックコードベクトルを上記残差出力 $r'(n)$ と上記短期予測逆フィルタ10からの残差出力 $r(n)$ との出力差 $e(n)$ のエネルギーが最小となるようにサーチすればよい。

【0028】上記ダイナミックコードブック1は、例えばサンプリング周波数を8 kHzとし、音声の1フレームが40サンプルから構成されているとしたとき、例えば128個(128通り)のダイナミックコードベクトルを有する。しかし、この第1の実施例では、従来例で用いられていた合成フィルタを用いていない。このため減算器8では、128個のコードベクトルに応じた残差出力 $r'(n)$ と上記短期予測逆フィルタ10からの残差出力 $r(n)$ を比較するだけである。つまり、40サンプルのデータのそれぞれについて128通りの比較演算を行うだけである。

【0029】すなわち、この第1の実施例は、128個のコードベクトルに応じた残差出力 $r'(n)$ と上記短期予測逆フィルタ10からの残差出力 $r(n)$ とを比較し、その出力差 $e(n)$ のエネルギーを最小とするように、端子12から最適なダイナミックコードベクトルを最適インデックス J_{opt} としてサーチし、取り出すことによって、現在の入力音声信号を符号化する。

【0030】次に、上記短期予測逆フィルタ10の動作を図2によって説明する。この図2では、図1の入力端子9から入力される音声信号 $S(n)$ の1フレーム当たりのサンプル数を N_s としている。上述したように本実施例ではサンプリング周波数 f_s を8 kHzとしており、サンプル数 N_s は例えば40である。また、上記ダイナミックコードブック1のサイズ(選択の対象となるダイナミックコードベクトルの数)を N_c とする。本実施例では上述したように例えば128個(通り)である。また、分析の対象となる入力音声信号を $S(n)$ とする。ここで、 n は0以上 N_s 未満である。また、その時点(n の時点)での上記ダイナミックコードブック1の状態、すなわち取り出されるダイナミックコードベクトルを $J(n)$ とする。ここで、 n は0以上 N_c 未満である。そして、任意の時点での上記短期予測逆フィルタ10による残差波形を $r(n)$ とする。ここで、 n は0以上 N_s 未満である。

【0031】通常、音声のスペクトルは、フォルマントのたった図2のAに示すようなエンベロープとなる。このエンベロープに高速フーリエ変換(FFT)を施すと図2のBに示すようなスペクトルとなり、ピークとピークの間がピッチとなる。このスペクトルを上記短期予測逆フィルタ10に入力すると、エンベロープが縮小された図2のCに示すような波形の残差出力 $r(n)$ が得られる。すなわち、ピッチが少し残り、フォルマント合成に依存した冗長性が失われたホワイトノイズのような波

形である。

【0032】上記ダイナミックコードブック1には、上記図2のCに示された残差出力 $r(n)$ に非常に良く似た図2のDに示すようなダイナミックコードベクトル $J(n)$ が記憶されている。そこで、この第1の実施例は、上記減算器8で残差出力 $r(n)$ と上記ダイナミックコードベクトル $J(n)$ に応じた残差出力 $r'(n)$ とを比較演算している。

【0033】ここで、上記残差出力 $r(n)$ は1フレームのサンプル数 N 、に応じて N 、次元のベクトルとみる 10
ことができ、これを r とする。また、ダイナミックコ

$$|r - t c_J|^2$$

$$= |c_J|^2 (t - (r, c_J) / |c_J|^2)^2 + |r|^2 - (r, c_J)^2 / |c_J|^2$$

【0036】となる。ここで、 r は固定であるから $(r, c_J)^2 / |c_J|^2$ を最大にするインデックス J_{opt} を選択することによりピッチを固定できる。

【0037】次に、上述した従来のVSELP方式の音声コーダとこの第1の実施例との演算量の差について具体的に説明する。まず、VSELP方式の音声コーダの演算量について述べる。VSELP方式の音声コーダにおいては、上記合成フィルタ58を用いて音を合成している。このため、音を合成するためのフィルタリングの系列数 $z_j(n)$ は、

【0038】

【数2】

$$Z_j(n) = \sum_{i=0}^{\min(n, J-1)} r(i-J) h(n-i)$$

【0039】と表せる。ここで、 $h(n-i)$ は上記合成フィルタ58の $(n-i)$ サンプルでのインパルス応答である。

【0040】また、上記 $z_j(n)$ は、

$$z_j(n) = z_{j-1}(n-1) + r(-J) h(n)$$

ここで、 $1 \leq n \leq N-1$ 、 $z_j(0) = r(-J) h(n)$ である。

【0041】この $z_j(n)$ は上記VSELPでは 40×128 である。次に、上記VSELP方式の音声コーダにおいて、ダイナミックコードブック51から再合成した音のパワー G_j は、

【0042】

【数3】

*コードブック1のインデックス J の時点から N 、サンプルを取り出したものをベクトルとして c_J とする。この状態で、 r と c_J の距離(近似度)を求め、それが最短となるインデックス J_{opt} を持って、このフレームにおけるピッチを抽出できる。このとき c_J のゲインは、問わない。

【0034】すなわち、 $|r - t c_J|$ が最小、 $|r - t c_J|^2$ が最小のとき、このフレームにおけるピッチ J を抽出できるので、

【0035】

【数1】

$$G_J = \sum_{n=0}^{N-1} (b'_j(n))^2$$

20 【0043】となる。ここで、 $b'_j(n)$ は、上記ダイナミックコードブックから出たコードベクトルの出力 $b_j(n)$ が上記合成フィルタ $H(z)$ を通った後の出力である。また、このパワー G_j は、

【0044】

【数4】

$$G_J = E_{J-1} + \sum_{n=0}^{N_J-1} (b'_j(n))^2$$

【0045】とも表せる。この従来のVSELPでは 20×128 となる。

【0046】また、ダイナミックコードブックと元の音声の内積 C_j は、

【0047】

【数5】

$$C_J = \sum_{n=0}^{N-1} b'_j(n) p'(n)$$

【0048】となり、このVSELP方式の音声コーダでは 40×128 となる。また、合成フィルタのインパルスレスポンス $h(n)$ は、この合成フィルタが10次のIIRフィルタであり、20サンプル分を用いるので 10×20 となる。

40 【0049】以上、従来のVSELP方式の音声コーダの演算量に対する本実施例の音声コーダの演算量を次の表1に示す。

【0050】

【表1】

	VSELP	第1の実施例による方式
Z_j	40×128	0
G_j	20×128	1×128
C_j	40×128	40×128
$h(n)$	10×20	0
逆フィルタ	0	40×10

【0051】すなわち、この第1の実施例のは、合成フィルタを用いていないので、音声を合成するためのフィルタリングの系列数 $z_j(n)$ は0であり、合成フィルタのインパルスレスポンス $h(n)$ も0である。

【0052】また、音のパワー G_j は、上記VSELP方式の音声コードのそれがコードブックから音を合成したときのパワーであり計算が複雑であるのに対し、単にダイナミックコードブックのパワーである。これは、上述した c_j の大きさということになる。つまり、 $|c_j|^2$ となり、この第1の実施例では 1×128 となる。

【0053】また、ダイナミックコードブックと元の音声の内積 C_j は、上記 r と c_j の大きさの二乗の積となり、本実施例でも 40×128 となる。

【0054】但し、この第1の実施例では、短期予測逆フィルタ10を用いているため、このフィルタリングの系列数を 40×128 として求めている。従来のVSELP方式の音声コードでは、逆フィルタを用いていないので、0である。

【0055】したがって、VSELP方式の音声コードと第1の実施例の音声コードの演算量とを比較すると、13000対5648となり、本実施例の音声コードの演算量は従来のVSELP方式の音声コードのその半以下となった。

【0056】次に、第2の実施例となる音声コードについて説明する。図3は第2の実施例となる音声コードの概略構成を示す機能ブロック図である。図3において、この音声コードは、ダイナミックコードブック21と、第1の固定コードブック22及び第2の固定コードブック23の合わせて3つのコードブックを有する。

【0057】入力端子29には例えばサンプリング周波数 $f_s = 8 \text{ kHz}$ でサンプリングされ、図示しないA/D変換器によってデジタル信号に変換された音声信号 $S(n)$ が入力される。この音声信号 $S(n)$ は減算器30に供給される。この減算器30には、端子34で0入力された合成フィルタ33からの0入力応答も供給される。この0入力応答は0入力が入上り合成フィルタ33を通る際に新しいLPC係数により形成されたものである。そして、この減算器30は上記音声信号 $S(n)$ から0入力応答を減算し、その減算結果を $P(n)$ として短期予測逆フィルタ31に供給する。この短期予測逆フ

10 ilterタ31には入力端子32からフィルタ状態を0にクリアするクリアフィルタ信号が供給され、入力音声信号 $S(n)$ の1フレームをフィルタリングした後、フィルタの状態を0にクリアして初期化する。この短期予測逆フィルタ31からの出力を残差出力 $r(n)$ とする。この残差出力 $r(n)$ は、減算器28に供給される。

【0058】ここで、上記ダイナミックコードベクトルは、乗算器24で係数 β が乗算されたあと加算器27に供給される。また、上記第1の固定コードブック22のコードベクトルは、乗算器25で係数 γ_1 が乗算されたあと上記加算器27に供給される。さらに、第2の固定コードブック23のコードベクトルは、乗算器26で係数 γ_2 が乗算されたあと上記加算器27に供給される。そして、上記加算器27の出力は、残差出力 $r'(n)$ として上記減算器28に供給される。

【0059】この減算器28は、上記残差出力 $r(n)$ と上記残差出力 $r'(n)$ との差を出力差 $e(n)$ としてエネルギー計算部35に供給する。このエネルギー計算部35は、残差出力 $r'(n)$ と上記短期予測逆フィルタ31からの残差出力 $r(n)$ との出力差 $e(n)$ のエネルギーを計算する。そして、その出力差 $e(n)$ のエネルギーが最小となるように、この第2の実施例は端子36からダイナミックコードブック21のダイナミックコードベクトルのインデックス J_{opt} をサーチし、取り出すことによって現在の入力音声信号を符号化する。

【0060】この第2の実施例は、上記第1の実施例と同様に従来例で用いられていた合成フィルタを減算器の前段に使用していない。このため減算器28では、128個のコードベクトルに応じた残差出力 $r'(n)$ と上記短期予測逆フィルタ31からの残差出力 $r(n)$ を比較するだけである。つまり、40サンプルのデータのそれぞれについて128通りの比較演算が行われるだけである。

【0061】この第2の実施例は、上記第1の実施例が1フレームのフィルタリングを行った後、フィルタステートをそのままの状態にして次のフレームの処理を行う、すなわちフィルタステートをずっと保持し、繰り返し使うのに対して、1フレームのフィルタリングを行った後、フィルタステートを0にクリアしてから、次のフレームの処理を行うものである。

50 【0062】これは、例えば、この音声コードと共に通

信系に用いられる音声デコーダ側でのマッチングをとるために有効である。通常、エンコーダ側には生の音声が入力されるので問題はないが、デコーダ側では、元の生の音声を知らない状態で音声を再合成するわけである。このときに上記第1の実施例のようにフィルタステートをずっと保持したままでは不都合が生じる。すなわち、前のフレームの内部状態のフィルタ係数が新しくなった状態で、入力音声信号 $S(n)$ から0入力応答を減算しておかないと悪影響が出てくる。そして、0入力を合成フィルタ33に通し、元の音声信号から減算したものを

上記短期予測合成フィルタ31に入力している。
【0063】この第2の実施例は、前のフレームの影響を除去するために、0入力応答を入力音声から減算して、短期予測逆フィルタ31で残差 $r(n)$ を求めているため、減算結果 $P(n)$ を有効に使い、さらにフィルタの内部状態が量子化されたデータで更新されていくためデコーダ側と全く同じ状態が再現され、エンコーダとデコーダのマッチングがとりやすくなる。

【0064】ここで、この第2の実施例は、上記第1の実施例と同様に演算量を半分以上に抑えられる。演算量削減の詳細な説明については、省略する。

【0065】次に、本発明の第3の実施例となる音声コーダについて説明する。図4は第3の実施例となる音声コーダの概略構成を示す機能ブロック図である。図4において、この音声コーダは、ダイナミックコードブック41と、第1の固定コードブック42及び第2の固定コードブック43の合わせて3つのコードブックを有する。

【0066】入力端子48には例えばサンプリング周波数 $f_s = 8\text{ kHz}$ でサンプリングされ、図示しないA/D変換器によってデジタル信号に変換された音声信号 $S(n)$ が入力される。この音声信号 $S(n)$ は短期予測逆フィルタ49に供給され、残差出力 $r(n)$ となる。この残差出力 $r(n)$ は、減算器50に供給される。

【0067】ここで、上記ダイナミックコードベクトルは、乗算器44で係数 β が乗算されたあと加算器47に供給される。また、上記第1の固定コードブック42のコードベクトルは、乗算器45で係数 γ_1 が乗算されたあと上記加算器47に供給される。さらに、第2の固定コードブック43のコードベクトルは、乗算器46で係数 γ_2 が乗算されたあと上記加算器47に供給される。そして、上記加算器47の出力は、残差出力 $r'(n)$ として上記減算器50及び合成フィルタ52に供給される。

【0068】上記減算器50は、上記残差出力 $r(n)$ と上記残差出力 $r'(n)$ との差を出力差 $e(n)$ としてエネルギー計算部51に供給する。このエネルギー計算部51は、上記出力差 $e(n)$ のエネルギーを計算する。そして、その出力差 $e(n)$ のエネルギーを最小とするように、この第3の実施例は、例えば端子55から

ダイナミックコードブック41のダイナミックコードベクトルのインデックス J_{opt} をサーチし、取り出す。

【0069】次に、このインデックス J_{opt} のコードベクトルと該コードベクトルの近傍のコードベクトルとを合成フィルタ52に通すことにより、音声 $S'(n)$ が得られる。この音声 $S'(n)$ は、減算器53に供給される。この減算器53には、上記入力音声 $S(n)$ も供給されている。この減算器53の出力は、出力差 $E(n)$ としてエネルギー計算部54に供給される。このエネルギー計算部54は、上記出力差 $E(n)$ のエネルギーを計算する。そして、この第3の実施例は、出力差 $E(n)$ のエネルギーを最小とするように、端子55からダイナミックコードブック41の最適ダイナミックコードベクトルの最適インデックス J'_{opt} をサーチし、その最適インデックス J'_{opt} を取り出すことによって現在の入力音声信号を符号化する。

【0070】以上、この第3の実施例は、上記短期予測逆フィルタ49からの残差出力 $r(n)$ とダイナミックコードベクトルに応じた残差出力 $r'(n)$ との出力差 $e(n)$ のエネルギーを最小とするようなダイナミックコードベクトルのインデックス J_{opt} を直接的にサーチ（第1のサーチ工程）し、このインデックス J_{opt} のコードベクトルとその近傍のコードベクトルとを合成フィルタ52に供給して音声 $S'(n)$ を合成し、この音声 $S'(n)$ と入力音声 $S(n)$ との出力差 $E(n)$ のエネルギーが最小となるようなダイナミックコードブックのダイナミックコードベクトルのインデックス J'_{opt} を間接的にサーチ（第2のサーチ工程）し、取り出すことによって現在の入力音声信号を符号化している。

【0071】このため、この第3の実施例は、第1のサーチ工程でラフにインデックスサーチを行い、第2のサーチ工程でその近傍をシビアにサーチすることができ、したがって、演算量を減らしても正確なコードベクトルのインデックスをサーチできる。

【0072】

【発明の効果】本発明に係る音声符号化方法は、過去の音声信号との相関を利用して現在の入力音声信号を分析し符号化する音声符号化方法において、検出工程が現在の入力音声信号を分析する工程の分析出力との距離が最短となる過去の音声信号の分析出力を検出し、符号化工程が上記検出工程の検出出力を基に現在の入力音声信号を符号化するので、演算量を従来よりも半減できる。

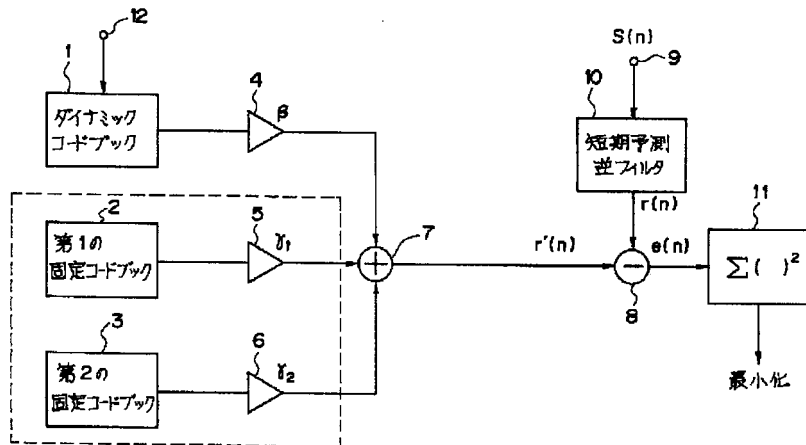
【0073】また、他の発明に係る音声符号化方法は、過去の音声信号の分析出力が複数のコードベクトルとして記憶されたコードブックをサーチして現在の入力音声信号との相関を利用して符号化を行う符号化方法であって、サーチ工程が現在の入力音声信号を分析する工程の分析出力との距離が最短となる上記コードブック内のコードベクトルを直接サーチし、符号化工程が上記サーチ工程で得られたコードベクトルのインデックスを用いて

現在の入力音声信号を符号化するので、演算量を従来よりも半減できる。

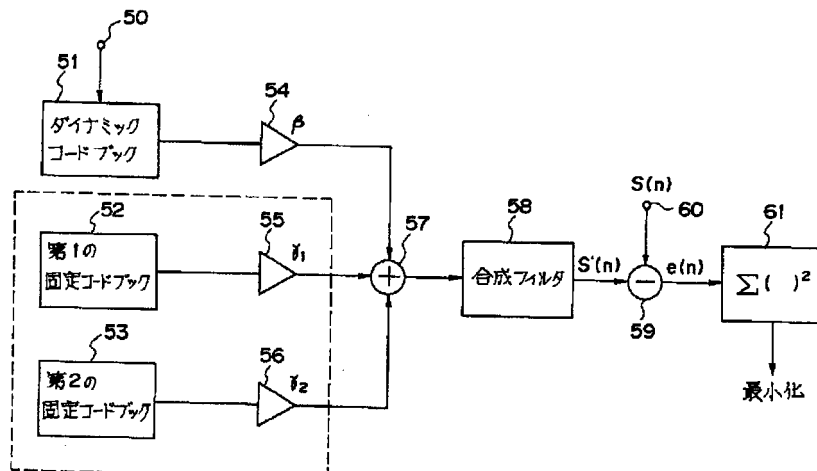
【0074】さらに、他の発明に係る音声符号化方法は、過去の音声信号の分析出力が複数のコードベクトルとして記憶されたコードブックをサーチして現在の入力音声信号との相関を利用して符号化を行う符号化方法であって、第1のサーチ工程が現在の入力音声信号を分析する工程の分析出力との距離が最短となる上記コードブック内のコードベクトルを直接サーチし、第2のサーチ工程が上記第1のサーチ工程で得られたコードベクトルを含め、該コードベクトルの近傍のコードベクトルの中から入力音声信号との相関が最適となるようなコードベクトルを間接サーチし、符号化工程が上記第2のサーチ工程で得られたコードベクトルのインデックスを用いて現在の入力音声信号を符号化するので、演算量を従来よりも減らせる。

*

【図1】



【図5】



*【図面の簡単な説明】

【図1】第1の実施例の音声コードの概略構成を示すブロック図である。

【図2】第1の実施例に用いられる短期予測逆フィルタの特性図である。

【図3】第2の実施例の音声コードの概略構成を示すブロック図である。

【図4】第3の実施例の音声コードの概略構成を示すブロック図である。

10 【図5】VSELPを用いた音声コードの機能ブロック図である。

【符号の説明】

1・・・ダイナミックコードブック

10・・・短期予測逆フィルタ

11・・・エネルギー計算部

Figure 1 consists of three vertically stacked plots labeled A, B, and C.

- Plot A:** A graph with the vertical axis labeled '音声レベル' (voice level) and an upward arrow, and the horizontal axis labeled '時間' (time) with a rightward arrow. It shows a continuous, smooth curve representing the speech level over time, with three distinct peaks of decreasing amplitude.
- Plot B:** A graph showing a sampled speech signal. The signal is a high-frequency oscillation whose amplitude envelope follows the curve in Plot A. A horizontal double-headed arrow between two peaks of the envelope is labeled 'ピッチ' (pitch).
- Plot C:** A graph showing a sampled signal $r(n)$ on the vertical axis. The horizontal axis is marked with $n=0$ and $n=N_s$. The signal is a high-frequency oscillation between $n=0$ and $n=N_s$.
- Plot D:** A graph showing a sampled signal $J(n)$ on the vertical axis. The horizontal axis is marked with $n=0$ and $n=N_c$. The signal is a high-frequency oscillation between $n=0$ and $n=N_c$.

Block diagram of a signal processing system (likely for speech synthesis or coding):

- Dynamic Codebook (21):** Receives input 36. Its output is multiplied by gain β (24).
- Fixed Codebooks (22, 23):** Labeled "第1の固定コードブック" and "第2の固定コードブック". Their outputs are multiplied by gains δ_1 (25) and δ_2 (26) respectively.
- Summing Node (27):** Sums the outputs from the dynamic codebook and the two fixed codebooks.
- Prediction Inverse Filter (31):** Labeled "短期予測逆フィルタ". It receives the sum $r'(n)$ (27) and feedback $P(n)$ (30).
- Subtractor (28):** Receives the output of the prediction inverse filter and the input signal $S(n)$ (29). Its output is $r(n)$ (28).
- Synthesis Filter (33):** Labeled "合成フィルタ". It receives $r(n)$ and produces output 34.
- Error Calculation (35):** A block labeled $\sum ()^2$ that calculates the squared error. It receives $r(n)$ and provides feedback $P(n)$ (30) to the prediction inverse filter.
- Minimum Error Output:** The final output of the error calculation block, labeled "最小化".

【図4】

